

## Durham Research Online

---

### Deposited in DRO:

03 September 2014

### Version of attached file:

Accepted Version

### Peer-review status of attached file:

Peer-reviewed

### Citation for published item:

Balasubramanian, Vijay and Berkooz, Micha and Ross, Simon F. and Simón, Joan (2014) 'Black holes, entanglement and random matrices.', *Classical and quantum gravity*, 31 (18). p. 185009.

### Further information on publisher's website:

<http://dx.doi.org/10.1088/0264-9381/31/18/185009>

### Publisher's copyright statement:

© 2014 IOP Publishing Ltd. This is an author-created, un-copyedited version of an article accepted for publication in *Classical and Quantum Gravity*. IOP Publishing Ltd is not responsible for any errors or omissions in this version of the manuscript or any version derived from it. The Version of Record is available online at <http://dx.doi.org/10.1088/0264-9381/31/18/185009>.

### Additional information:

## Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

# Black Holes, Entanglement and Random Matrices

Vijay Balasubramanian<sup>a,b,1</sup>, Micha Berkooz<sup>c,2</sup>, Simon F. Ross<sup>d,3</sup>, Joan Simón<sup>e,4</sup>

<sup>a</sup>*David Rittenhouse Laboratories, University of Pennsylvania,  
209 S 33<sup>rd</sup> Street, Philadelphia, PA 19104, USA*

<sup>b</sup>*Initiative for the Theoretical Sciences, CUNY Graduate Center,  
365 Fifth Avenue, New York, NY 10016, USA*

<sup>c</sup>*Department of Particle Physics and Astrophysics, Weizmann Institute of Science,  
Rehovot 76100, Israel*

<sup>d</sup>*Centre for Particle Theory, Department of Mathematical Sciences,  
Durham University, South Road, Durham DH1 3LE, UK*

<sup>e</sup>*School of Mathematics and Maxwell Institute for Mathematical Sciences,  
University of Edinburgh, King's Buildings, Edinburgh EH9 3JZ, UK*

## Abstract

We provide evidence that strong quantum entanglement between Hilbert spaces does not generically create semiclassical wormholes between the corresponding geometric regions in the context of the AdS/CFT correspondence. We propose a description of low-energy gravity probes as random operators on the space of black hole states. We use this description to compute correlators between the entangled systems, and argue that a wormhole can only exist if correlations are large. Conversely, we also argue that large correlations can exist in the manifest absence of a Lorentzian wormhole. Thus the strength of the entanglement cannot generically diagnose spacetime connectedness, without information on the spectral properties of the probing operators. Our random matrix picture of probes also provides suggestive insights into the problem of “seeing behind a horizon”.

## 1 Introduction

Famously, the eternal black hole in AdS space, understood as a wormhole between two asymptotically AdS regions, is dual to two non-interacting conformal field theories in a thermally entangled state [1]. Generalizing this idea, it has been proposed that spacetime connectedness in AdS space is related to quantum entanglement in the dual field theory [2], and, further, that entanglement is equivalent to the existence of wormholes in spacetime

---

<sup>1</sup>email: vijay@physics.upenn.edu

<sup>2</sup>email: micha.berkooz@weizmann.ac.il

<sup>3</sup>email: S.F.Ross@durham.ac.uk

<sup>4</sup>email: j.simon@ed.ac.uk

[3].<sup>5</sup> There is some controversy over how general the relation between entanglement and wormholes is. It was conjectured in [3] that entanglement should be *identified* with the existence of a wormhole (ER=EPR). However Marolf and Polchinski [8] used the eigenvalue thermalization hypothesis (ETH) to argue that the local correlations in a typical entangled state are weak, and hence should not correspond to a semiclassical wormhole in the bulk. Shenker and Stanford [9, 10] found examples of special states corresponding to long semiclassical wormholes, where the local correlations are weak but a smooth wormhole exists. In this paper we will examine this question using a model based on describing low-energy probes in the bulk as random matrices acting on the space of states of a black hole. In this random matrix model we will find a suppression of correlations in the typical state (unlike for the thermofield double state), in agreement with [8], and argue that this implies that these typical states do not have a semiclassical wormhole interpretation.

We consider a Hilbert space  $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_R$ , where  $\mathcal{H}_{L,R}$  are identical and dynamically independent factors. A particular entangled state in this Hilbert space is the thermofield double state

$$|\psi_\beta\rangle = \frac{1}{Z(\beta)} \sum_i e^{-\beta E_i/2} |i\rangle_L \otimes |i\rangle_R, \quad (1)$$

where  $E_i$  is the energy of the eigenstate labeled by  $i$  in the  $L$  and  $R$  Hilbert spaces, and  $Z(\beta)$  normalizes the state. Tracing over  $\mathcal{H}_L$  gives a thermal density matrix in  $\mathcal{H}_R$ . This state can be thought of as a purification of the thermal density matrix and is identified in AdS/CFT with the eternal black hole in the bulk. One Hilbert space factor is associated to each of the two asymptotic boundaries, and the entropy of the reduced density matrix on  $\mathcal{H}_R$  is identified with the area of the horizon, that is, with the minimal cross-sectional area of the Einstein-Rosen bridge (wormhole) between the asymptotic regions. Thus, the entropy of the reduced density matrix diagnoses the size of the wormhole. Furthermore, the entanglement in (1) gives rise to finite “two-sided” correlation functions  $\langle \mathcal{O}_L \mathcal{O}_R \rangle$  between operators supported on  $\mathcal{H}_{L,R}$  respectively. In AdS space, this correlator is computed in a suitable approximation from spacelike geodesics which link the two boundaries of spacetime through the wormhole.

Now consider some more general entangled state on  $\mathcal{H}_L \otimes \mathcal{H}_R$  which reduces to the thermal density matrix when one factor is traced over. Consider CFT operators dual to supergravity fields for which a probe approximation in the bulk is appropriate (that is, where the effects of back-reaction of this operator insertion can be neglected; we will assume in particular that the insertion changes the energy by an amount  $\Delta E \lesssim T$ ) and where the operator dimension  $\Delta \gg 1$  so that the geodesic approximation [11] to bulk correlators is reliable.<sup>6</sup> Then the existence of a wormhole would imply that the two-point function between insertions of this operator in the two entangled copies of the field theory,  $\langle \mathcal{O}_L \mathcal{O}_R \rangle$ , will be given by a geodesic passing through the wormhole, and hence should be of roughly the same order as the two-point function in a single copy of the CFT,  $\langle \mathcal{O}_R \mathcal{O}_R \rangle$ ; we would not expect it to be suppressed by any factor of the dimension of the Hilbert space. Thus, if the two-point function  $\langle \mathcal{O}_L \mathcal{O}_R \rangle$  is exponentially suppressed relative to  $\langle \mathcal{O}_R \mathcal{O}_R \rangle$  by factors

---

<sup>5</sup>In a related development based on the Ryu-Takayanagi expression for the entanglement entropy in field theory in terms of minimal surfaces in AdS space [4], the areas of bulk surfaces have been reconstructed from a “differential entropy” measured from the entanglement structure of the field theory state [5, 6]. This may be related to the proposal in [7] to reconstruct the bulk spacetime from the entanglement structure of the field theory state using tensor network techniques from condensed matter physics.

<sup>6</sup>We will also include objects like D-branes in this class. These two limits can be simultaneously realized by making the temperature  $T$  or the typical energy of the entangled states sufficiently large.

involving the dimension of the space in which the entanglement occurs (that is, the entropy of the reduced density matrix in one copy of the Hilbert space), we take this as evidence that the state does not correspond to a semi-classical wormhole.

In Sec. 2, we will argue that the low-energy gravity probes we’re interested in can be approximated as random matrices acting on the space of states of a black hole, because of the inability of low-energy supergravity modes to probe typical states of black holes [12, 13, 14]. This leads to our key methodological innovation: the matrix elements of a given operator are modelled by a matrix drawn at random from some suitable ensemble. Given that the operator  $\mathcal{O}$  can be modelled by a random matrix, we can approximate the correlator  $\langle \mathcal{O}_L \mathcal{O}_R \rangle$  by an average over the operator ensemble. This gives us a simple calculation to test the validity of the wormhole interpretation for a given state.

Modelling the probe by a random matrix is a stronger statement than the usual idea that excitations of a thermodynamic system thermalize to local thermal equilibrium; in an ordinary system like a lump of coal, we have some spatial resolution for our probes, so the excitation does not act totally randomly in that it will initially excite some local subset of degrees of freedom. Probing black holes is supposed to be more difficult, in that once the excitation has fallen into the black hole it is no longer localized to a subset of the degrees of freedom but acts truly randomly.

In section 3, we treat the gravity probes as matrices acting on the states with energy in some narrow range  $(E - \Delta, E + \Delta)$ , defining a microcanonical Hilbert space  $\mathcal{H}_E$ , of dimension  $d_E = e^{S(E)}$ . We use a uniform random matrix ensemble on  $\mathcal{H}_E$  to model the operators. This allows us to do computations in states

$$|\psi_U\rangle = \frac{1}{d_E} \sum_{i,j \in \mathcal{H}_E} U_{ij} |i\rangle_L \otimes |j\rangle_R \quad (2)$$

where  $U_{ij}$  is some unitary matrix. For all of these states, tracing over one Hilbert space factor gives the maximally mixed density matrix

$$\rho_E = \frac{1}{d_E} \mathbb{I}_E. \quad (3)$$

with entanglement entropy  $S(E)$ . One of these states is the microcanonical analogue of (1):  $|\psi\rangle = (1/Z(\beta)) \sum_i e^{-\beta E_i} |i\rangle_L \otimes |i\rangle_R \sim e^{-S(E)/2} \sum_i |i\rangle_L \otimes |i\rangle_R$ . The question is whether the “two-sided” two-point functions  $\langle \psi_U | \mathcal{O}_L \mathcal{O}_R | \psi_U \rangle$  could be interpreted in terms of a wormhole in AdS. We find that it is suppressed relative to (1) by a factor of  $1/d_E \sim e^{-S}$ , where  $S$  is the entropy of the reduced density matrix. By contrast, if we pick  $U$  to be the identity (i.e. the state is the microcanonical analogue of the thermally entangled state (1)) this suppression vanishes. Thus, the two-sided correlators in generic states  $|\psi_U\rangle$  with the same entanglement entropy as the thermal state do not have the structure expected to allow a dual description in terms of a classical wormhole.

The uniform random matrix ensemble of operators on  $\mathcal{H}_E$  provides a basic approximation to the properties of low-energy supergravity modes, but it does not fully capture the physics of the correlation functions of these operators, notably their time dependence. In section 4 we show that by introducing more general matrix ensembles where we allow transitions between states of different energies, we can reproduce the exponential decay of correlators determined by the quasi-normal modes in the AdS description, supporting our proposal that physics of probes of complex gravitational states can be understood in terms of the dynamics of random matrices. The ensembles introduced in Sec. 4 also allow

us to make closer contact with the work of [8] which uses the eigenvalue thermalization hypothesis.

In Sec. 5 we turn to an important subtlety: at low temperatures, below the Hawking-Page transition, the thermofield double state (1) is no longer dual to a black hole, and is rather described in the bulk as two copies of thermal AdS. (More accurately, the latter saddle point of the gravitational path integral dominates the over the black hole saddle point.) We will see via analytic continuation from the Euclidean theory that the two-sided correlators in this state remain large despite the absence of a Lorentzian wormhole. This emphasizes that these correlators can only provide a necessary, and not a sufficient, condition for the presence of a wormhole. As we discuss, our picture of black hole probes remains consistent with this sort of transition if, below the transition, operators no longer act like random matrices. This seems natural because this is now an ordinary weakly interacting gas, where the probe will excite the gas locally.

Section 6 discusses further implications of a random matrix description of operators. In particular such a description gives correlators a robust analytic structure in the complex time plane irrespective of the state or ensemble which they probe. We speculate that this could lead to a universal “behind the horizon” continuation for the observables described by our random matrix approximation. We also note that random operators give rise to a Wick-like scheme of contractions, independent of the usual one in terms of creation and annihilation operators, and suggest that this gives rise to a modified perturbative scheme at the horizon.

**Relation to other work:** In [8], Marolf and Polchinski invoked the eigenvalue thermalization hypothesis to consider the behaviour of the “two-sided” two-point functions  $\langle \mathcal{O}_L \mathcal{O}_R \rangle$  in typical entangled states on  $\mathcal{H}_L \otimes \mathcal{H}_R$ . They found the same suppression by  $e^{-S}$ . The ideas underlying our calculation are similar, but we approximate gravity operators as random matrices, rather than assuming eigenvalue thermalization. One might regard our computations as an effective model reproducing the physics that would arise if eigenvalues thermalize. The idea that the amount of quantum entanglement is not enough to guarantee the existence of semiclassical wormholes was also advocated in [15]. These authors dispute the interpretation of the eternal BH as a thermally entangled state [1], but our results support the latter idea.

In [10], Shenker and Stanford provided examples of states which have weak local correlations but which do have a smooth classical dual. (Other work on deformations of the thermofield double state includes [16].) However, as noted in [10], these states are not generic and it seems difficult to construct a smooth geometry with the correct properties to give the dual of a generic state with weak correlations. We therefore interpret the exponentially suppressed two-point functions that we find as a signal of the absence of a geometric connection in the dual. The wormholes discussed in [10] have a natural interpretation as states close to the thermofield double (1) in a sense that could be made precise using the random matrix approach.

For the thermofield double state (1), two-sided correlations are large at  $t = 0$  but decay as time increases. In [17, 18] it has been speculated that this decay is related to the increasing complexity of the state, with the length  $\ell$  of the wormhole in units of the AdS scale  $L$  being given by

$$\frac{\ell}{L} \sim \frac{\mathcal{C}}{S}, \quad (4)$$

where  $\mathcal{C}$  is the computational complexity of the state and  $S$  is the entanglement entropy.

This would correspond to the two-sided correlations scaling as

$$\langle \mathcal{O}_L \mathcal{O}_R \rangle \sim e^{-\Delta \frac{\mathcal{C}}{S}}, \quad (5)$$

where  $\Delta$  is the conformal dimension of  $\mathcal{O}$ . Our results are qualitatively consistent with this picture, as the states  $|\psi_U\rangle$  are more complex than (1), and have weaker correlations. However, if we apply this interpretation to our results (equivalently, those of [8]), it would correspond to assigning a complexity  $\mathcal{C} \sim S^2$  to our state  $|\psi_U\rangle$ , which is much smaller than the maximal complexity  $\mathcal{C}_{max} \sim e^S$ . It will be interesting to understand this relation in more detail.

There is some tension between our random matrix picture and the idea that the black hole microstates can be modelled by smooth geometries or extended D-brane configurations [19, 20], as one would expect probes of such local objects to excite them locally as with ordinary thermodynamic systems. The resolution may lie in unknown aspects of the dynamics of such microstate geometries. One could also envision some splitting of the Hilbert space into subspaces such that the operator acts as random operators within each subspace.

## 2 Random operators

Our key idea is that operators corresponding to low-energy gravity probes act randomly on the space of states associated with a black hole. We first want to clarify the extent of our claim. Consider a typical pure CFT state  $|\psi\rangle$  dual to a large black hole.<sup>7</sup> We act on this state with some CFT operator  $\mathcal{O}$  dual to a supergravity mode. There is considerable evidence that local operators in the CFT which do not significantly change the total energy are insensitive to the particular details of the state  $|\psi\rangle$  [12, 13, 14]. Therefore we would expect the transitions between such states produced by the action of this operator to be essentially random.

However, such a random matrix model is not expected to capture all of the physics of the operator insertion. At short times, the change in the state  $|\psi\rangle$  retains some structure which reflects which operator  $\mathcal{O}$  we acted with. This is indicated for example by the two-point correlation function of an operator with its conjugate, which at very short distances is dominated by the singular part of the OPE. We refer to this as the structured part of the operator. But at longer times (presumably compared to some thermalization scale of the state), or when acting on the degrees of freedom associated with a black hole, the ability to know the final state after acting with  $\mathcal{O}$  is exponentially small. We call this part of the operator's action the unstructured part of the operator, and it is this that we want to model by a random matrix.

The distinction can be easily understood in the dual bulk perspective in the geodesic approximation. The structured part of a correlation function is associated with geodesics which remain far from the black hole horizon, while the unstructured part involves geodesics which pass close to the horizon, or disconnected paths where one particle is absorbed by the black hole and another one is emitted in the Hawking radiation. That this is a useful diagnostic can be seen by considering situations in which the operator will remain structured for a long time, for example by considering operators with large angular momentum on the global AdS black hole, and noting such cases are reflected in the existence of geodesics which stay outside the black hole for a long time. Thus, even though modeling the gravity

---

<sup>7</sup>That is, where the black hole is the dominant contribution to the dual bulk description; this excludes low-temperature states where the bulk description is thermal AdS, see comments on this in section 5.

operators as random operators does not capture all aspects of the operators, it should capture precisely the part that we need. This is because the issue of whether there is a semiclassical wormhole or not is an issue of what happens behind the horizon.

Thus, *low energy gravity modes act, to a good approximation, as random matrices on the states of the black hole.* In other words, gravity modes encode the minimum possible information about the actual microstates of black holes. We will thus treat the matrix elements of a given operator  $\mathcal{O}$  as drawn at random from a suitable matrix ensemble. We will define the ensembles of interest in the next two sections.

In this paper, we use this random matrix description to provide a criterion for the existence of a wormhole in the dual gravitational description of a typical entangled state. As argued in the introduction, we would expect states with a wormhole interpretation to have the property that the two-sided correlator  $\langle \mathcal{O}_L \mathcal{O}_R \rangle$  is of the same order as the one-sided correlator  $\langle \mathcal{O}_R \mathcal{O}_R \rangle$ . Given the random matrix description, we can approximate these correlators by considering the average over the matrix ensemble that the operators are drawn from. For simple matrix ensembles, it is then easy to test this criterion in generic entangled states.

The use of an ensemble average is also supported by the fact that in the geodesic approximation, the bulk two-point function calculation is largely insensitive to the details of the individual operator being considered, depending only on its conformal dimension. One might however still be concerned that the average could be suppressed relative to the value for a particular operator by phase cancellation. But in section 3 we will see that when we consider the state (1) the average remains of the same order as for a given operator, which provides some evidence against this possibility. We will also see that standard deviations in the ensemble averages are exponentially small.

### 3 Entanglement vs wormholes: fixed energy

In this section, we consider entangled states which involve energy eigenstates lying in a narrow range of energies, and restrict attention to operators acting within this energy range. That is, we work with states belonging to the subspace  $\mathcal{H}_E \otimes \mathcal{H}_E \subset \mathcal{H}_L \otimes \mathcal{H}_R$ , where  $\mathcal{H}_E$  contains exact energy eigenstates  $|i\rangle \in \mathcal{H}_E$  with eigenvalues  $E_i \in [E - \Delta, E + \Delta]$ , and we assume that there is a large density of states at these energies. Entangled states can be written in this energy basis as

$$|\psi_c\rangle = \sum_{i,j} c_{ij} |i, j\rangle \quad \text{with} \quad \sum_{i,j} |c_{ij}|^2 = 1 \quad (6)$$

where  $|i, j\rangle = |i\rangle_L \otimes |j\rangle_R$ .

A particularly interesting subset of quantum pure states in  $\mathcal{H}_E \otimes \mathcal{H}_E$  is defined by the property that tracing over  $\mathcal{H}_L$  gives rise to the microcanonical ensemble, i.e. the maximally mixed density matrix

$$\rho_E = \frac{1}{d_E} \sum_{i \in \mathcal{H}_E} |i\rangle \langle i| = \frac{\mathbb{I}_E}{d_E}, \quad (7)$$

where  $d_E = e^{S(E)}$  is the dimension of  $\mathcal{H}_E$ . We will denote this set of states in the Hilbert

space by  $\mathcal{H}_U$ .<sup>8</sup> It includes all the states in  $\mathcal{H}_E \otimes \mathcal{H}_E$  of the form

$$|\psi_U\rangle = e^{-S/2} \sum_{i,j} U_{ij} |i,j\rangle \quad (8)$$

for any unitary matrix  $U \in U(d_E)$ . These states have the same amount of entanglement as the state

$$|\psi_{\text{micro}}\rangle = |\psi_{\mathbb{I}}\rangle = e^{-S/2} \sum_{i \in \mathcal{H}_E} |i,i\rangle \quad (9)$$

which provides the standard purification of the single sided microcanonical density matrix.<sup>9</sup> Restricting to states with the same reduced density matrix is useful because it allows us to see clearly that the details and not just the overall amount of entanglement between the two Hilbert spaces plays a key role in the emergence of a smooth wormhole. We will study how single sided and two sided correlators behave in various  $|\psi_c\rangle$  and  $|\psi_U\rangle$ .

In subsection 3.1 we define the operator ensemble we consider, which is just the ensemble of gaussian random matrices in  $\mathcal{H}_E$ . In section 3.2 we evaluate single sided correlators in the various states. In section 3.3 we compute the two sided correlators for various states, and compare them with the single sided correlators on the same states. In section 3.4 we compute the standard deviations of the various correlators.

### 3.1 Operator averaging & random matrices

To model operators that act within  $\mathcal{H}_E$  we will assume that operator matrix elements are drawn from the simplest Gaussian distribution

$$\mathcal{F}_r = \frac{1}{Z_r^M} dM_{ij} dM_{ij}^* e^{-\gamma \text{tr}(M M^\dagger)}, \quad (10)$$

where we denote the matrix elements  $\langle i|\mathcal{O}|j\rangle$  as  $M_{ij} \forall |i\rangle, |j\rangle \in \mathcal{H}_E$ , and  $Z_r^M$  is a normalization factor, chosen so that  $\int \mathcal{F}_r = 1$ . We will refer to this as the *restricted* operator ensemble, as it applies to operators which are restricted to act within  $\mathcal{H}_E$ . This ensemble is assumed to be universal for all operators acting within this Hilbert space. We consider calculations where we take the ensemble average within the ensemble of operators (10), keeping the state  $|\psi_c\rangle$  fixed.

This choice of ensemble is motivated by simplicity: it is the gaussian matrix ensemble invariant under unitary transformations of  $\mathcal{H}_E$  which depends only on the dimensionality of  $\mathcal{H}_E$ .<sup>10</sup> The gaussian assumption amounts to a sort of free-field approximation for the operators, as in the ensemble (10) the only non-trivial connected correlation function is the two point function

$$\mathbf{E}(M_{ij}^* M_{kl}) = \frac{1}{\gamma} \delta_{ik} \delta_{jl}. \quad (11)$$

(We will use the notation  $\mathbf{E}$  to stress that we are taking expectation values in our operator ensemble.) Thus, when we insert operators in higher-point correlation functions, the ensemble expectation values will be determined by a Wick-like pair-wise contraction of insertions of operators using (11) (after appropriately summing over the indices). This should

<sup>8</sup>Note that  $\mathcal{H}_U$  is not a subspace of the Hilbert space as a vector space, as the requirement that the reduced density matrix is (7) is not a linear constraint on the Hilbert space.

<sup>9</sup>This is the microcanonical ensemble equivalent of the thermofield double state (1).

<sup>10</sup>This is true assuming that we do not impose any restrictions of hermiticity or unitarity on  $\mathcal{O}$ .



be related to the free field approximation in the bulk spacetime, which is valid at leading order in  $N$  in the large  $N$  limit of the CFT. We will comment further on corrections from including higher order polynomials in the ensemble measure and the perturbation theory around this free-field behaviour in section 6.2.

The ensemble involves a single parameter  $\gamma$ , which depends on the energy  $E$  used to define  $\mathcal{H}_E$ . The scaling of  $\gamma$  with  $E$  can be determined by considering the “inclusive cross section” - starting from a given initial state  $|i_0\rangle \in \mathcal{H}_E$ , acting on it by the operator  $\mathcal{O}$ , and ending in all possible states (in  $\mathcal{H}_E$ ). This gives

$$\mathbf{E} \left( \sum_{k \in \mathcal{H}_E} |\langle k | \mathcal{O} | i_0 \rangle|^2 \right) = \frac{e^{S(E)}}{\gamma}, \quad (12)$$

so requiring that the inclusive cross-section is finite implies

$$\gamma = \hat{\gamma} e^{S(E)}, \quad (13)$$

that is, the parameter  $\gamma$  should scale like the dimension  $d_E = e^{S(E)}$  of  $\mathcal{H}_E$ .

### 3.2 Single sided correlators

Since our criterion is based on a comparison between single sided and double sided two-point functions (to avoid issues related to operator normalization), we will begin by computing the operator ensemble average of single sided correlators using (10).

We will begin with  $|\psi_U\rangle$ . Since we are computing single sided correlators we can reduce to the single sided density matrix (7) first:

$$\begin{aligned} \langle \psi_U | \mathcal{O}_R^\dagger(t) \mathcal{O}_R(0) | \psi_U \rangle &= \text{tr}_{\mathcal{H}_R} \left( \rho_E \mathcal{O}_R^\dagger(t) \mathcal{O}_R(0) \right) \\ &= e^{-S} \sum_{i \in \mathcal{H}_E, n \in \mathcal{H}_R} e^{i(E_i - E_n)t} |\langle n | \mathcal{O}_R | i \rangle|^2 \end{aligned} \quad (14)$$

Under the assumption that  $\mathcal{O}$  acts within  $\mathcal{H}_E$  (which allows us to sum only over  $|n\rangle \in \mathcal{H}_E$ ), we can easily estimate the size of this two-point function at  $t = 0$  to be of order one

$$\mathbf{E}(\langle \psi_U | \mathcal{O}_R^\dagger(0) \mathcal{O}_R(0) | \psi_U \rangle) = e^{-S} \sum_{i \in \mathcal{H}_E, n \in \mathcal{H}_E} \mathbf{E}(|\langle n | \mathcal{O}_R | i \rangle|^2) = e^{-S} \frac{e^{2S}}{\gamma} = \frac{1}{\hat{\gamma}}. \quad (15)$$

Our choice of scaling for  $\gamma$  thus has the nice consequence that correlators allowing a semi-classical gravitational interpretation are order one, i.e. they do not scale in the dimension of the microcanonical ensemble  $d_E = e^S$ . Computing the full time dependent operator ensemble average two-point function is just as easy and it is given by

$$\mathbf{E}(\langle \psi_U | \mathcal{O}_R^\dagger(t) \mathcal{O}_R(0) | \psi_U \rangle) = \frac{e^{-S}}{\gamma} \sum_{i,n} e^{i(E_i - E_n)t} = \frac{e^{-2S}}{\hat{\gamma}} |\text{tr}_{\mathcal{H}_E}(W(t))|^2 \quad (16)$$

where  $W(t) = e^{-iHt}$ . The time dependence is generated by the slight variations in the energy in the range from  $E - \Delta$  to  $E + \Delta$ , which produces only a very slow time variation of the correlator, on the times scales where we resolve these small energy differences. We will have a detailed discussion of time dependence in section 4, where we consider an ensemble of

operators which includes larger energy transitions, and we can model the bulk gravitational time dependence successfully.

We could consider more general states  $|\psi_c\rangle$  in a similar way, but one easily sees that the correlators will be qualitatively the same as the ones above for any state whose reduced density matrix is localised in  $\mathcal{H}_E$ . In fact they are qualitatively the same even for pure states  $|\psi_V\rangle = \sum_i V_i |i\rangle \in \mathcal{H}_E$ ,

$$\mathbf{E}(\langle\psi_V|\mathcal{O}_R^\dagger(t)\mathcal{O}_R(0)|\psi_V\rangle) = \frac{e^{-S}}{\hat{\gamma}} \langle\psi_V|W(t)^*|\psi_V\rangle \text{tr}_{\mathcal{H}_E}(W(t)). \quad (17)$$

This is a straightforward consequence of our starting assumption that the operators act randomly on states in  $\mathcal{H}_E$ . If the operators have simple spectral properties - i.e., they are drawn randomly from a simple matrix ensemble - then single sided correlators can't even distinguish whether the state is a density matrix or a pure state. For two sided correlators, the story will be different, but for a single sided observer, measuring any finite number of correlators will not distinguish between a pure state, the thermal state, or any other mixed state with the same macroscopic quantum numbers.

### 3.3 Two-sided correlators

We can now compute two-sided correlators, i.e. those involving operators acting on both Hilbert spaces  $\mathcal{H}_L$  and  $\mathcal{H}_R$ . We first review the analytic continuation between the one-sided and two-sided correlators in a thermal ensemble, which teaches us that we want to consider two-sided correlators where the operator on  $\mathcal{H}_L$  is time reversed.<sup>11</sup> We then discuss the calculation of these two-sided correlators in the operator ensemble average, and compare to the one-sided case.

#### 3.3.1 Mapping of operators

In the thermal ensemble, we can define operators acting on  $\mathcal{H}_L$  in terms of the action on  $\mathcal{H}_R$  by the formula (see for example [27])

$$\begin{aligned} \text{tr}(\rho_\beta \mathcal{O}_{R,\alpha_1}(t_1) \dots \mathcal{O}_{R,\alpha_n}(t_n) \mathcal{O}_{R,\beta_l}(t'_l + i\beta/2) \dots \mathcal{O}_{R,\beta_1}(t'_1 + i\beta/2)) = \\ = \langle \Psi | \mathcal{O}_{R,\alpha_1}(t_1) \dots \mathcal{O}_{R,\alpha_n}(t_n) \tilde{\mathcal{O}}_{L,\beta_1}(t'_1) \dots \tilde{\mathcal{O}}_{L,\beta_l}(t'_l) | \Psi \rangle \end{aligned} \quad (18)$$

Here  $t_1, \dots, t_n$  and  $t'_1, \dots, t'_n$  are separately time ordered from earliest to latest times, and the  $\alpha_i$  and  $\beta_i$  simply label the operators. Also,  $\rho_\beta$  is the canonical thermal density matrix on  $\mathcal{H}_R$ , while

$$|\Psi\rangle = \frac{1}{\sqrt{Z(\beta)}} \sum_{\alpha} e^{-\beta E_\alpha/2} |\alpha, \alpha\rangle \quad (19)$$

is the thermofield double state. For the particular case of two point functions, inserting the thermal density matrix on the left hand side and the state  $|\Psi\rangle$  on the right hand side gives,

$$\sum_{i,j} e^{-\beta E_i} \langle i | \mathcal{O}_{R,\alpha_1}(t) | j \rangle \langle j | \mathcal{O}_{R,\beta_1}(t' + i\beta/2) | i \rangle = \sum_{i,j} e^{-\frac{\beta}{2}(E_i + E_j)} \langle i | \mathcal{O}_{R,\alpha_1}(t) | j \rangle \langle i | \tilde{\mathcal{O}}_{L,\beta_1}(t') | j \rangle \quad (20)$$

<sup>11</sup>In general, the states in  $\mathcal{H}_L$  are the CPT conjugates of the ones in  $\mathcal{H}_R$ . Since we are primarily interested in discussing the time dependence in our correlators, we stressed the time reversal.

where on the left hand side we inserted a complete set of states  $|j\rangle$ . The two terms in the exponential on the right hand side come from the bra and the ket in the correlator. Next, we can write the imaginary time translation of  $O_{R,\beta_1}$  in the second term on the left hand side as  $O_{R,\beta_1}(t' + i\beta/2) = e^{i(i\beta/2)H} O_{R,\beta_1}(t') e^{-i(i\beta/2)H}$ . Acting on the state vectors on either side of  $O_{R,\beta_1}(t')$  this produces exponential factors that now match between the left and right hand sides. So we can conclude that  $\langle j|O_{R,\beta_1}(t')|i\rangle = \langle i|\tilde{O}_{L,\beta_1}(t')|j\rangle$ . In other words  $\tilde{O}_{L,\beta_1}(t') = O_{R,\beta_1}(t')^T$ . There is one final subtlety – in the conventional thermofield double description and in the eternal black hole, global time is defined to run backwards in the second (L) copy. To be consistent with this convention we should flip the direction of time for the  $O_L$  operators. If we choose  $t = 0$ ,  $t' = 0$ , to form the initial Cauchy surface and flip the time direction in  $\mathcal{H}_L$  to align time with global time on both sides, we finally have

$$\mathcal{O}_L(t) = \tilde{\mathcal{O}}_L(-t) = \mathcal{O}_R(-t)^T. \quad (21)$$

In a general state  $|\psi_c\rangle$ , we will therefore consider correlators between an operator  $\mathcal{O}_R$  acting on  $\mathcal{H}_R$  and the operator  $\mathcal{O}_L$  acting on  $\mathcal{H}_L$  defined by (21). Comparing this two-sided correlator to the one-sided correlator  $\langle \mathcal{O}_R \mathcal{O}_R \rangle$  will give us our criterion for the existence of wormholes.

### 3.3.2 Two-sided correlators and semiclassical ER bridges

Given a pure state  $|\psi_U\rangle \in \mathcal{H}_L \otimes \mathcal{H}_R$ , the two point two sided correlator is then

$$\langle \psi_U | \mathcal{O}_R^\dagger(t) \mathcal{O}_L(0) | \psi_U \rangle = e^{-S(E)} \sum_{i,j,k,l} U_{ij} U_{kl}^* e^{i(E_l - E_j)t} \langle i | \mathcal{O}_R(0) | k \rangle \langle l | \mathcal{O}_R^\dagger(0) | j \rangle. \quad (22)$$

The ensemble average is

$$\begin{aligned} \mathbf{E}(\langle \psi_U | \mathcal{O}_R^\dagger(t) \mathcal{O}_L(0) | \psi_U \rangle) &= e^{-S(E)} \sum_{i,j,k,l} U_{ij} U_{kl}^* e^{i(E_l - E_j)t} \mathbf{E}(M_{jl}^* M_{ik}) \\ &= \frac{e^{-2S(E)}}{\hat{\gamma}} |\text{tr}(U W(t))|^2, \end{aligned} \quad (23)$$

where we used (11) and  $W(t) = e^{-itH}$ .

We would like to compare this expression with the single sided two-point function (16). It is enough to focus on  $t = 0$ . Since the trace is at most of order  $d_E = e^{S(E)}$ , the two sided correlator is bounded by the single sided one (15), and they are the same only when  $U \propto \mathbb{I}_E$ . We interpret this as saying that the wormhole connecting the two spaces is as large or as semiclassical as it can be when we have the standard purification of the microcanonical density matrix.

For most choices of  $U$ , the correlator will be much smaller. To determine the value for a typical  $U$ , we can consider now drawing  $U$  itself uniformly from the ensemble of random unitary matrices, corresponding to choosing a typical state in  $\mathcal{H}_U$ .

$$\mathbf{E}_U(\mathbf{E}(\langle \psi_U | \mathcal{O}_R^\dagger(0) \mathcal{O}_L(0) | \psi_U \rangle)) = \frac{1}{d_E^2 \hat{\gamma}} \sum_{i,j} \int dU U_{ii} U_{jj}^* = \frac{1}{d_E^3 \hat{\gamma}} \sum_{i,j} \delta_{ij} = \frac{1}{d_E^2 \hat{\gamma}}, \quad (24)$$

where  $\mathbf{E}_U$  stands for the average over the uniform distribution of unitary matrices.<sup>12</sup> The

<sup>12</sup>In a previous version of the paper, we obtained an estimate scaling as  $1/d_E$  by diagonalizing  $U$  and assuming that the eigenvalues  $e^{i\theta_i}$  are drawn uniformly from the circle. This is not correct, as the Jacobian factor in passing from the integral over  $U_{ij}$  to the eigenvalues introduces an eigenvalue repulsion. This repulsion produces the additional suppression in the average correlator found here.

average value of the two-sided correlation function is smaller than the single-sided one by a factor of  $1/d_E^2 = e^{-2S}$ .

We will see below that the standard deviation in the average over operators is larger than this average, scaling as  $e^{-S}$ , so the correct estimate for the two-sided correlator for a particular operator is smaller than the single-sided one by a factor of  $1/d_E = e^{-S}$ . We conclude that these typical states cannot have a smooth wormhole description in the gravitational dual. The same conclusion was reached in [8] by appealing to the eigenvalue thermalization hypothesis [21, 22]. Our approach based on random matrices gives a different, more computationally tractable perspective on the result.

It would be interesting to study the behaviour of this two-sided correlation function for small deformations of  $U = \mathbb{I}_E$ , and compare it to the changes in the length of the wormhole, which were recently studied by Shenker & Stanford [9, 10]. We leave this for future work.

### 3.4 Standard deviations

Deviations within the ensemble give us a measure of the departure of the correlator for a specific operator from the ensemble average considered above. We would expect these deviations to be small when the dimension of the Hilbert space is large.

The standard deviations in our operator ensemble are

$$\sigma_{\mathcal{O},RR}^2(U, t) = \mathbf{E}(|\langle \psi_U | O_R^\dagger O_R | \psi_U \rangle|^2) - |\mathbf{E}(\langle \psi_U | O_R^\dagger O_R | \psi_U \rangle)|^2, \quad (25)$$

$$\sigma_{\mathcal{O},RL}^2(U, t) = \mathbf{E}(|\langle \psi_U | O_R^\dagger O_L | \psi_U \rangle|^2) - |\mathbf{E}(\langle \psi_U | O_R^\dagger O_L | \psi_U \rangle)|^2. \quad (26)$$

The one-sided quantity (25) will obviously be independent of the unitary matrix  $U$ . It is

$$\sigma_{\mathcal{O},RR}^2(U, t) = e^{-2S(E)} \sum_{i,j,k,l} e^{i(E_i - E_j)t} e^{-i(E_k - E_l)t} \left( \mathbf{E}(M_{ji}^* M_{ji} M_{lk} M_{lk}^*) - \mathbf{E}(M_{ji}^* M_{ji}) \mathbf{E}(M_{lk}^* M_{lk}) \right). \quad (27)$$

The only non-trivial contribution comes from the contractions  $\mathbf{E}(M_{ji}^* M_{lk}) \mathbf{E}(M_{ji} M_{lk}^*)$ . These give

$$\sigma_{\mathcal{O},RR}^2(U, t) = \frac{e^{-2S(E)}}{\hat{\gamma}^2}. \quad (28)$$

Note that the index contractions are such that the phases cancel, so the variation is time-independent. The computation of (26) is very similar. In this case,

$$\sigma_{\mathcal{O},RL}^2(U, t) = e^{-2S(E)} \sum_{i,j,k,l} U_{ij} U_{kl}^* e^{i(E_l - E_j)t} U_{i'j'}^* U_{k'l'} e^{-i(E_{l'} - E_{j'})t} \left( \mathbf{E}(M_{jl}^* M_{ik} M_{j'l'} M_{i'k'}^*) - \mathbf{E}(M_{jl}^* M_{ik}) \mathbf{E}(M_{i'k'}^* M_{j'l'}) \right). \quad (29)$$

Its non-trivial contribution comes from the same contractions as before, and the index contraction is such that the factors of  $U$  cancel out in addition to the time dependence, giving

$$\sigma_{\mathcal{O},RL}^2(U, t) = \frac{e^{-2S(E)}}{\hat{\gamma}^2}. \quad (30)$$

We can compare these standard deviations to the average size of the correlators,

$$\begin{aligned} \frac{\sigma_{\mathcal{O},RR}^2(\mathbf{U}, t)}{|\mathbf{E}(|\langle \psi_{\mathbf{U}} | O_R^\dagger O_R | \psi_{\mathbf{U}} \rangle)|^2} &= \frac{e^{2S}}{|\text{tr} W(t)|^4}, \\ \frac{\sigma_{\mathcal{O},RL}^2(\mathbf{U}, t)}{|\mathbf{E}(|\langle \psi_{\mathbf{U}} | O_R^\dagger O_L | \psi_{\mathbf{U}} \rangle)|^2} &= \frac{e^{2S}}{|\text{tr} U W(t)|^4}. \end{aligned} \quad (31)$$

For the single sided correlators, so long as  $t\Delta \ll 1$  (where  $\Delta$  is the energy spread in  $\mathcal{H}_E$ ) the standard deviation is small compared to the average, as expected. For the two-sided correlators, the answer depends on the state under consideration, i.e. the unitary matrix  $U$ . For the standard purification  $\mathbf{U} = \mathbb{I}$ , the one sided and two sided correlators are the same size. But for typical unitary matrices  $\text{tr} U \sim e^{S/2}$ , so already for  $t = 0$ , the standard deviation is larger than the correlator. Thus, for a typical state  $|\psi_{\mathbf{U}}\rangle$ , the overall size of the correlator for a particular operator should be estimated from the standard deviation; the typical value is thus smaller than the single-sided one by  $e^{-S}$ , and the value fluctuates from operator to operator, producing a smaller average value. This also seems problematic for attempts to interpret the correlations as due to a smooth semiclassical wormhole.

## 4 Operator ensemble including energy-changing transitions

In the previous section, we assumed that the operator changes the energy only by a small amount. This restriction may be too strong for some operators, so in this section we will provide a model for random operators when allowing for transitions between more disparate energy states. This will be useful for considering correlations in states corresponding to the canonical ensemble, and will allow us to model the bulk time-dependence associated with quasinormal modes of the black hole.

When we allow transitions that change the energy, not all states are on an equal footing and the distribution of matrix elements can change as a function of the initial and final state energies. The operator ensemble we consider to model this behaviour should then be more complicated. Following the philosophy of effective field theory, we write this matrix distribution as an expansion in the energy separation of the states involved in the transition, assuming the existence of an averaged energy  $E$  around which we work. The leading terms in the distribution of matrices are then

$$\mathcal{F}_g = \frac{1}{Z_g^M} \prod_{ij} dM_{ij} dM_{ij}^* e^{-\gamma(\text{tr}(MM^\dagger) - \alpha_1 \text{tr}([M, H]M^\dagger) + \alpha_2 \text{tr}([H, M][M^\dagger, H]) + \dots)}, \quad (32)$$

where  $Z_g^M$  ensures the normalisation condition  $\int \mathcal{F}_g = 1$  is satisfied. We will refer to this as the *energy-changing operator ensemble*.

In the spirit of effective field theory, the parameters  $\alpha_i$  will be determined by matching the properties of the correlators in the ensemble to those of correlators in a bulk black hole for some particular field. We will see that the resulting values guarantee convergence of the matrix integrals in  $\int \mathcal{F}_g$ . Also, a priori  $|i\rangle, |j\rangle$  run over all the states in the Hilbert space, but transitions between highly separated energies are suppressed in the ensemble by  $\alpha_2$ , and the size of the energy transitions for the values determined from the black hole will be small compared to the overall energy in the thermodynamic limit.

Since we retain in  $\mathcal{F}_g$  only terms quadratic in the matrix elements, we can rewrite (32) as

$$\mathcal{F}_g \propto \prod_{ij} dM_{ij} dM_{ij}^* e^{-\gamma \sum_{kl} \Delta_{kl} |M_{kl}|^2}, \quad (33)$$

with

$$\Delta_{kl} = 1 + \alpha_1(E_k - E_l) + \alpha_2(E_k - E_l)^2 + \dots \equiv P(\alpha_j, E_k - E_l), \quad (34)$$

where in the last step we introduced a general polynomial of the energy difference  $P(E_k - E_l)$  to emphasize that many of our statements can be extended to higher orders in  $E_k - E_l$ . For the most part we will truncate the expansion to second order, but when specific details will not matter, we will use the arbitrary polynomial  $P$ . As in subsection 3.1, the only non-trivial connected correlation function is

$$\mathbf{E}(M_{ij}^* M_{kl}) = \frac{1}{\gamma \Delta_{ij}} \delta_{ik} \delta_{jl}. \quad (35)$$

Let us first discuss the scaling of  $\gamma$  in this ensemble. As in (12), we will require the total “inclusive cross-section” for transitions between any states of energy  $E_i$  and  $E_j$  to be finite. We would now like to do this while preserving the symmetry of the transition amplitude  $M_{ij}$  in  $i$  and  $j$ . It is then convenient to choose  $\gamma$  as

$$\gamma = \hat{\gamma}(e^{S(E_i)} + e^{S(E_j)}). \quad (36)$$

This entails a minor modification of (32) that moves  $\gamma$  inside the trace<sup>13</sup>. There is some arbitrariness in our definition of the scaling of  $\gamma$ , but in the effective field theory approach that we are discussing, and when we allow transitions only between close by  $E_k$  and  $E_l$ , then other choices which keep the “inclusive cross section” finite amount to redefinitions of the  $\alpha$ s. We will use the choice above because it is computationally simple.

The “inclusive cross section” is then

$$\sum_{k \in \mathcal{H}_{E_j}} |\langle k | M | i_0 \rangle|^2 \sim \frac{e^{S(E_j)}}{\hat{\gamma}(e^{S(E_j)} + e^{S(E_{i_0})}) P(\alpha_l, E_j - E_{i_0})}, \quad (37)$$

which is finite for all  $|i_0\rangle$  and  $E_j$ . We should actually require the above expression to be finite when we integrate over  $E_k$ , but we will see that this happens automatically because of the  $\text{tr}([M, H]^2)$  terms.

Let us now determine the coupling constants  $\alpha_l$  and understand their physical relevance by matching observables. As said earlier, we will choose these by matching the correlators in the ensemble to those of a particular operator. That is, we now model a given operator  $\mathcal{O}$  as a matrix chosen at random from the ensemble (32) with ensemble parameters  $\alpha_l$  chosen to reproduce some of the expected structure of the operator  $\mathcal{O}$ . Consider the one sided two point function in the canonical ensemble,

$$\begin{aligned} \frac{1}{Z(\beta)} \mathbf{E} \left( \text{tr}(e^{-\beta H} \mathcal{O}_R^\dagger(t) \mathcal{O}_R(0)) \right) &= \frac{1}{Z(\beta)} \sum_{ik} e^{-\beta E_i} e^{i(E_i - E_k)t} \mathbf{E}(M_{ki}^* M_{ki}) \\ &= \frac{1}{Z(\beta)} \int dE_i dE_k \frac{e^{-\beta E_i + S(E_i) + S(E_k) + i(E_i - E_k)t}}{\hat{\gamma}(e^{S(E_i)} + e^{S(E_k)}) P(\alpha_j, E_k - E_i)}. \end{aligned} \quad (38)$$

Note that this integral is convergent, as large  $E_i$  are cut-off by the  $-\beta E_i$  term and for large  $E_k$ , the numerator and denominator  $e^{S(E_k)}$  cancel, leaving us with terms which are

---

<sup>13</sup>Strictly speaking  $\hat{\gamma}$  should also be inside the trace. However, this will have no consequences for our conclusions.

oscillating in  $E_k$ . Changing variables to  $E_i$  and the energy difference  $\Delta = E_k - E_i$ , the integral becomes approximately

$$\int d\Delta \frac{e^{-it\Delta}}{\hat{\gamma}(1 + e^{-\beta\Delta}) P(\alpha_l, \Delta)}. \quad (39)$$

In the last formula we cancelled the  $E_i$  integral against  $Z(\beta)$  (picking up corrections which scale like  $1/E$  where  $E$  is the averaged energy, as we show in the appendix).<sup>14</sup>

For large  $t$  we deform the contour of  $\Delta$  to the upper half plane and pick up the leading pole of the denominator. Denoting it by  $\Delta_0$ ,  $\text{Im}(\Delta_0) > 0$ , the correlator decays exponentially at long times as

$$e^{i\Delta_0 t}. \quad (40)$$

In the bulk spacetime, the exponential decay of a particular correlator is determined by the lowest quasi-normal mode of the corresponding field; i.e.,  $\Delta_0$  corresponds to the complex frequency of the lowest quasi-normal mode of the bulk black hole. Thus, given a bulk field, we choose the coefficients  $\alpha_l$  in the definition of the ensemble to reproduce the quasi normal modes of the bulk field. In the quadratic approximation we focus on,  $P(\alpha_l, \Delta) = 1 + \alpha_1 \Delta + \alpha_2 \Delta^2$ , and  $\alpha_{1,2}$  are determined by requiring that this polynomial has a zero at the complex frequency of the lowest quasi-normal mode. If we retained higher-order terms in the polynomial, these could be determined by requiring further zeros match the complex frequencies of higher quasi-normal modes.

Additionally,  $\alpha_1$  encodes the commutator  $[\mathcal{O}_R^\dagger, \mathcal{O}_R]$  since

$$\begin{aligned} \mathbf{E}(\text{tr}(\frac{e^{-\beta H}}{Z(\beta)} [\mathcal{O}_R^\dagger, \mathcal{O}_R])) &= \frac{1}{Z(\beta)} \sum_{i,k} e^{-\beta E_i} \mathbf{E}(M_{ki}^* M_{ki} - M_{ik} M_{ik}^*) \\ &= \frac{1}{Z(\beta)} \sum_{i,k} \frac{e^{-\beta E_i}}{\gamma} \left( \frac{1}{\Delta_{ki}} - \frac{1}{\Delta_{ik}} \right) \end{aligned} \quad (41)$$

In the restricted operator ensemble we truncated to  $E_l = E_k$  and dropped the commutator terms in the brackets. This can be realized as a limit of our more general operator ensemble by rescaling  $\alpha_2 \rightarrow \infty$ ,  $\alpha_1 \rightarrow 0$  with an appropriate rescaling of  $\hat{\gamma}$ . Thus, the restricted operator ensemble will be a good approximation for times short compared to the lowest quasi normal frequency of the bulk field.

## 4.1 Single sided correlators

Having determined the scaling behaviour of the parameters characterizing our operator ensemble (32), we can now calculate ensemble averages of correlation functions. In this subsection we consider the single-sided correlators, which depend only on the reduced density matrix obtained from the state.

We calculated the one-sided two-point function for the thermal density matrix in (38), and fixed the parameters  $\alpha_l$  by requiring that it have an exponential decay in time which

<sup>14</sup>To do so we expanded  $S(E_k) = S(E_i - \Delta E)$  to 1st order in  $\Delta E$ . This approximation is valid since the  $\Delta E^2$  term in this expansion is multiplied by a quantity which is of order  $1/E$  but  $\Delta E$  is finite and does not scale with the total energy, so this is small. Similarly the  $\alpha$ 's depend on the energy in which we evaluate them - ie, they hide an  $E_i$  dependence. However, taking this dependence into account introduces correction terms proportional to  $\partial_E \alpha$  and  $\partial_E \beta(E)$ . These terms again scale as  $1/E$  and hence are negligible.

corresponds qualitatively to the expected exponential decay of correlations due to the quasi-normal modes in the bulk black hole. Reproducing this exponential decay (that is, having a corresponding short time scale in the operator ensemble) was one of the primary motivations for generalizing from the restricted operator ensemble.

We can also easily calculate the one-sided two-point function in the microcanonical ensemble of states specified by an energy  $E$ ,  $\text{tr} \left( \rho_E \mathcal{O}_1^\dagger(t) \mathcal{O}_1(0) \right)$ . The difference from the previous section is that we allow intermediate states at energies different from  $E$ . This gives

$$\int dE_k \frac{e^{S(E_k) + i(E_i - E_k)t}}{\hat{\gamma}(e^{S(E_i)} + e^{S(E_k)})(1 + \alpha_1(E_k - E_i) + \alpha_2(E_k - E_i)^2)}, \quad (42)$$

which is the same integral as we had in (38) (after switching to  $\Delta E = E_i - E_k$  and expanding for small  $\Delta E$ ). The fact that we get the same result from the canonical and microcanonical ensemble of states (up to  $1/E$  corrections) is what we expect from standard thermodynamic considerations.

We can easily extend the calculation of the one-sided two-point function to an arbitrary pure state in  $\mathcal{H}_L \otimes \mathcal{H}_R$

$$|\psi_c\rangle = \sum_{i,j \in \mathcal{H}} c_{ij} |i, j\rangle, \quad (43)$$

where  $\sum_{i,j \in \mathcal{H}} |c_{ij}|^2 = 1$ . This includes as particular cases:

- Microcanonical entangled state:  $c_{ij} = e^{-S(E)/2} \delta_{ij}$  for  $|i\rangle, |j\rangle \in \mathcal{H}_E$ , zero otherwise
- Thermofield double:  $c_{ij} = \delta_{ij} e^{-\beta E_i/2}$
- Unitary twisting of the microcanonical entangled state (which reduces to the microcanonical ensemble on tracing over one side):  $c_{ij} = e^{-S(E)/2} U_{ij}$  for  $|i\rangle, |j\rangle \in \mathcal{H}_E$ , zero otherwise
- Unitary twisting of the thermofield double:  $c_{ij} = U_{ij} e^{-\beta E_i/2}$
- Pure state:  $c_{ij} = V_i^* V_j$ .

One sided correlators only depend on the reduced density matrix

$$\rho_c = \text{tr}_{\mathcal{H}_L} |\psi_c\rangle \langle \psi_c| = \sum_{ij} (c^\dagger c)_{ji} |i\rangle \langle j|. \quad (44)$$

The one sided two-point function equals

$$\text{tr} \left( \rho_c \mathcal{O}_R^\dagger(t) \mathcal{O}_R(0) \right) = \sum_{i,j,m} (c^\dagger c)_{ji} e^{it(E_j - E_m)} M_{mi} M_{mj}^*, \quad (45)$$

thus its operator ensemble average in the energy-changing ensemble equals

$$\mathbf{E} \left( \text{tr} \left( \rho_c \mathcal{O}_R^\dagger(t) \mathcal{O}_R(0) \right) \right) = \sum_{i,m} (c^\dagger c)_{ii} \frac{e^{it(E_i - E_m)}}{\gamma \Delta_{mi}}. \quad (46)$$

In this case the correlator is

$$\int dE_k \sum_i (c^\dagger c)_{ii} \frac{e^{S(E_k) + i(E_i - E_k)t}}{\hat{\gamma}(e^{S(E_i)} + e^{S(E_k)})(1 + \alpha_1(E_k - E_i) + \alpha_2(E_k - E_i)^2)} \quad (47)$$



Recall that our normalization condition is  $\sum_i (cc^\dagger)_{ii} = 1$ . The particular cases considered above are sharply peaked around some energy  $E$  (at least when the ensembles are stable) so in all these cases the correlator is qualitatively the same. In particular this is true for a pure state  $c_{ij} = V_i^* V_j$ . This is just a consequence of the fact that we can divide our states into very fine energy slices, and then the ensemble is invariant under arbitrary unitary transformations in each of the slices. Thus the single-sided probes are not sensitive to any detailed information about the state.

#### 4.1.1 Standard deviations

To study how reliable the ensemble averages (46) are, we compute the variances of single sided two-point correlators in our operator ensemble,

$$\sigma_{cRR}^2 = \mathbf{E} \left( \left| \text{tr} \left( \rho_c \mathcal{O}_R^\dagger(t) \mathcal{O}_R(0) \right) \right|^2 \right) - \left| \mathbf{E} \left( \text{tr} \left( \rho_c \mathcal{O}_R^\dagger(t) \mathcal{O}_R(0) \right) \right) \right|^2. \quad (48)$$

The first term is

$$\mathbf{E} \left( \left[ \text{tr} \left( \rho_c \mathcal{O}_R^\dagger(t) \mathcal{O}_R(0) \right) \right]^2 \right) = \sum_{i,j,\dots m,n} (c^\dagger c)_{ij} (c^\dagger c)_{kl}^* e^{it(E_j - E_m)} e^{-it(E_l - E_n)} \langle M_{mi} M_{mj}^* M_{nk}^* M_{nl} \rangle. \quad (49)$$

The only contribution to the difference (48) comes from contractions in which the operators  $\mathcal{O}_R$  are contracted between traces. Thus, we are only left with the contraction  $\langle M_{mi} M_{nk}^* \rangle \langle M_{mj} M_{nl}^* \rangle$ . Thus the variance is

$$\sigma_{cRR}^2 = \sum_{i,j,m} \frac{|(c^\dagger c)_{ij}|^2}{\gamma^2} \frac{1}{\Delta_{mi} \Delta_{mj}} = \sum_{i,j} |(c^\dagger c)_{ij}|^2 \int dE \frac{e^{S(E)}}{\gamma^2 P(E - E_i) P(E - E_j)}, \quad (50)$$

where we took the continuum limit in  $E_m = E$  in the last step.

Plugging in the expressions for  $c_{ij}$  in the canonical, microcanonical and pure state we obtain the following scaling behaviour for  $\sigma_{cRR}^2$ : Canonical, Microcanonical  $\sim e^{-2S(E)}$ , Pure state  $\sim e^{-S(E)}$ .

## 4.2 Two sided correlators

In this section we compute the two-sided correlation function in the general pure state  $|\psi_c\rangle$  defined in (43). Applying the map of operators from  $\mathcal{H}_R$  to  $\mathcal{H}_L$  discussed in section 3.3.1, the two sided two-point function in an arbitrary pure state  $|\psi_c\rangle$  is

$$\langle \psi_c | \mathcal{O}_R^\dagger(t) \mathcal{O}_L(0) | \psi_c \rangle = \sum_{i,j,k,l} c_{ij} c_{kl}^* e^{it(E_l - E_j)} M_{ik} M_{jl}^*. \quad (51)$$

Thus, its operator ensemble average is

$$\mathbf{E}(\langle \psi_c | \mathcal{O}_R^\dagger(t) \mathcal{O}_L(0) | \psi_c \rangle) = \sum_{i,k} c_{ii} c_{kk}^* \frac{e^{i(E_k - E_i)t}}{\gamma \Delta_{ik}}. \quad (52)$$

This should be compared with the single sided correlator for the same  $|\psi_c\rangle$ , in equation (46).

For the special cases corresponding to the canonical and microcanonical ensemble where  $c_{ij} \propto \delta_{ij}$ , we can see that the correlators have the same form as the single-sided one. For the

unitary transformed state with  $c_{ij} = U_{ij}e^{-\beta E_j/2}$ , where  $U$  is an arbitrary unitary matrix, as in the restricted operator ensemble, the two sided correlator will depend on  $U$ . We define the "coherence density" as

$$F(E) = e^{-S(E)} \sum_{i \in \mathcal{H}} U_{ii}, \quad (53)$$

then the two sided correlator is

$$\mathbf{E}(\langle \psi_c | \mathcal{O}_R^\dagger(t) \mathcal{O}_L(0) | \psi_c \rangle) = \frac{1}{Z} \int dE_i dE_k F(E_i) F^*(E_k) \frac{e^{-\beta E_i/2 - \beta E_k/2 + S(E_i) + S(E_k)}}{e^{S(E_i)} + e^{S(E_k)}} \frac{e^{i(E_k - E_i)t}}{\hat{\gamma} \Delta_{ik}(E_i, E_k)} \quad (54)$$

For  $U_{ij} = \delta_{ij}$ ,  $F$  is of order one, and this two-sided correlator is similar to the one-sided correlator. For generic  $U_{ij}$ ,  $F$  is exponentially suppressed, and as in our previous discussion in the microcanonical operator ensemble, the two-sided correlator is exponentially suppressed relative to the single-sided correlator, so we conclude that the states do not have a smooth wormhole description in the dual.

Again, it would be interesting to compare cases where  $F(E)$  is order one to the discussion of semi-classical throats of greater length in [9]. For the canonical ensemble, we could take  $c_{ij} = e^{-\beta E_i/4 - \beta E_j/4} V_{ij}$  where  $V$  is close to the identity, and mixes only states very close in energy, say less than  $\delta$ . In this case  $F(E) = e^{-S(E) + \mathcal{O}(\beta\delta)} \sum_{i \in \mathcal{H}_E} V_{ii} = 1 + \mathcal{O}(\beta\delta)$ .

#### 4.2.1 Standard deviations

To study how reliable the ensemble averages (52) are, we compute the standard deviations of two sided two-point correlators in our operator ensemble,

$$\sigma_{cRL}^2 = \mathbf{E} \left( \left| \langle \psi_c | \mathcal{O}_R^\dagger(t) \mathcal{O}_L(0) | \psi_c \rangle \right|^2 \right) - \left| \mathbf{E} \left( \langle \psi_c | \mathcal{O}_R^\dagger(t) \mathcal{O}_L(0) | \psi_c \rangle \right) \right|^2. \quad (55)$$

The calculation proceeds as in section 4.1.1. In particular, the origin of the contractions giving rise to a non-trivial contribution is entirely the same. The variance equals

$$\sigma_{cRL}^2 = \sum_{i,j,k,l} \frac{|c_{ij}|^2 |c_{kl}|^2}{\gamma^2} \frac{1}{\Delta_{ik} \Delta_{jl}}. \quad (56)$$

For the  $c_{ij}$  corresponding to both the standard ensemble states and their generalized versions with a unitary matrix, this scales as  $e^{-2S}$ , as in the discussion in the restricted operator ensemble. For the standard microcanonical and canonical states this then implies the standard deviation is small compared to the average value, while for the general states involving  $U_{ij}$ , it is of the same order as the average value.

Thus the results for two-sided correlators in this more general energy-changing operator ensemble are qualitatively the same as in the simpler restricted ensemble.

### 4.3 Relation to the eigenvalue thermalization hypothesis

We have introduced these random matrix ensembles as a way to model the behaviour of operator correlation functions in a complicated thermal system (namely the black hole states). We would like to compare this to a previous description of such correlators, the

eigenvalue thermalization hypothesis (ETH) [21, 22]. The ETH assumes<sup>15</sup>

$$\langle k | \mathcal{O}_L | i \rangle = \mathcal{A}(\bar{E}) \delta_{ki} + e^{-S(\bar{E})/2} f^A(E_\alpha, E_\beta) M_{\alpha\beta}^L \quad (57)$$

$$\langle k | \mathcal{O}_R | i \rangle = \mathcal{B}(\bar{E}) \delta_{ki} + e^{-S(\bar{E})/2} f^B(E_\alpha, E_\beta) M_{\alpha\beta}^R \quad (58)$$

$$M_{ki}^L M_{mn}^R = \delta_{km} \delta_{in} \sigma^{LR}(E_\alpha, E_\beta) + \text{erratic} \quad (59)$$

In our analysis, we have assumed that the one point functions  $\mathcal{A}$  and  $\mathcal{B}$  are approximately vanishing; this is typical for non-conserved quantities in the thermodynamic limit, but our model could easily be extended to include such terms. The main point of comparison is then the second part of the ETH matrix elements.

Our previous microcanonical operator ensemble can be compared to the ETH matrix elements with  $E_\alpha = E_\beta$ . These are matrices whose only non-trivial connected Green's function is a two point function + erratic. This corresponds precisely to the behaviour of our microcanonical ensemble in (11); the fluctuations of individual elements of the ensemble about the average will produce the erratic term. (By erratic one means a random noise which averages to zero.) So our restricted operator ensemble is one example of an ensemble that reproduces the behaviour assumed in ETH, for operators which make transitions between nearby energies.

In the energy-changing ensembles, our correlators have again the structure assumed in ETH, but we have introduced a different description of the dependence on the energy differences. In ETH, this is characterized by the functions  $f^A$ ,  $f^B$  and  $\sigma^{LR}$ , which are assumed to be smooth functions of  $E_\alpha, E_\beta$ , but are otherwise arbitrary. Our model goes a step further, by providing a simple “effective action” description of this dependence for energies  $E_\alpha, E_\beta$  near some fixed energy. This effective action approach allows us to fix the energy dependence to reproduce aspects of the black hole behaviour like the quasinormal mode. This provides a nice way to build more specific models of the operator matrix elements; our approach could easily be extended to include features of interest for other specific applications.

## 5 Structured operators and the confined phase

In this section, we first note that there are cases in which a thermal density matrix does not correspond to an Einstein-Rosen wormhole, but the two sided correlators are still large. This is a simple General Relativity statement, which has been somewhat neglected in the discussion of connection between entanglement and geometry. We then conjecture that this may be related to the fact that the probing gravity operators are not random on the thermal states, but rather they are “structured” in the sense of section 2. In general we want to speculate that the random nature of low-energy gravity probes is essential for understanding how entanglement is related to geometry.

It is well-known that if we consider a field theory on the sphere, there is a Hawking-Page phase transition at a finite temperature (of order the radius of the sphere) where the thermal AdS and black hole saddle-points exchange dominance. Thus, for a thermofield double state at sufficiently low temperature, the dual bulk description does not involve a large wormhole; rather, it is two copies of thermal AdS, with the thermal graviton gas on the

---

<sup>15</sup>We note that the separation of the operator matrix elements into a smooth  $f$  and  $M^L$  seems somewhat artificial, since both of them are really matrices that depend of  $\alpha$  and  $\beta$ , and hence it's not clear where to draw the line between them.

two spaces in an entangled state. Nonetheless, the two-sided two-point functions remain of the same order as the one-sided ones. For the field theory, this is automatic: since we are in a thermal state, the one- and two-sided correlators are related by analytic continuation as in section 3.3.1. In the bulk, these large two-sided correlators can be understood through the existence of a connection between the two boundaries in the Euclidean section. (So we can still interpret the two-sided correlator in terms of a geodesic length,  $\langle \mathcal{O}_L \mathcal{O}_R \rangle \sim e^{-\Delta_\ell}$ , where  $\ell$  is now the length of a geodesic in the Euclidean instanton linking the two boundaries.) But there is no Lorentzian wormhole linking the two AdS spaces. From the Lorentzian point of view, these correlators are large because of the entanglement between the bulk modes in the two thermal AdS spaces.

The von Neumann entropy of the reduced density matrix on one side below this transition is order one, so one might think that the absence of a smooth wormhole in this case is associated with the small amount of entanglement between the two sides. But if we work in the microcanonical ensemble, we can increase the energy to reach regimes where the entropy is large where the disconnected saddle is still the dominant bulk solution.<sup>16</sup>

In our analysis, this phase is qualitatively different from the black hole phase because the gravity mode operators acting on thermal AdS will not behave like random operators; probing a thermal gas in AdS, these operators are sensitive to the differences between different states. They are sparse operators which are related in a specific way to the states that make up the ensemble, i.e, they are “structured” with the respect to the states. Thus, it is possible that what makes the difference between the geometric connection in the black hole case and its absence here is that in the former we had large correlations for random operators, whereas here we have them for some operators which are structured with respect to the states.

We now present a simple model to make the point that gravity modes will be structured operators in this phase. Consider a Hilbert space made out of many harmonic oscillators with frequencies  $w_1, \dots, w_k$ , with total energy  $H = \sum n_i w_i$ . In the AdS/CFT picture, these are the energies of the single particle states in the bulk. We will denote the states by  $|\vec{n}\rangle$  where  $\vec{n}$  is a vector of integers of length  $k$ , which tells us the particle number for each energy level  $w_i$ . When we need to distinguish a specific particle - particle  $l$  - from the rest we will denote the state as  $|\vec{n}, n_l\rangle$ . The operators that we use to probe the geometry are then the corresponding raising or lowering operators. That is, we have at our disposal operators  $T_i$ ,  $i = 1..k$ , and our model for their action is

$$\langle \vec{n}, n_i | T_i | \vec{m}, m_i \rangle = \delta_{\vec{n}, \vec{m}} (\delta_{n_i+1, m_i} + \delta_{n_i, m_i+1}) \quad (60)$$

where the  $i$ 'th index is dropped from  $\vec{n}$  and  $\vec{m}$ . These operators are clearly not generating random transitions between the energy eigenstates; instead, the operator matrix elements are sparse matrices.

Consider now the state

$$\frac{1}{\sqrt{Z}} \sum_{\vec{n}} e^{-\beta E_{\vec{n}}/2} |\vec{n}, \vec{n}\rangle. \quad (61)$$

This could be thought of as either 1) a model of the canonical ensemble in the full CFT for temperatures below the Hawking-Page transition, or 2) a canonical description of the gas of gravity modes when pushed to  $T \gg 1$ , in the regime where it is locally stable. We can

---

<sup>16</sup>In the microcanonical ensemble, the thermal gas of gravitons is the dominant phase up to a Hagedorn transition, at energies set by the 't Hooft coupling in the field theory; thus the entropy will grow up to some power of the 't Hooft coupling [23].

compare the single and two sided correlators of the structured operators  $T_i$  in this state. A simple calculation gives

$$\langle T_1^{(1)}(t) T_1^{(1)}(0) \rangle = e^{-iw_1 t} + e^{iw_1 t - \beta w_1}, \quad (62)$$

while

$$\langle T_1^{(2)}(t) T_1^{(1)}(0) \rangle = (e^{-iw_1 t} + e^{iw_1 t}) e^{-\beta w_1/2}. \quad (63)$$

The interpretation is simplest for the case that  $w_1 \beta \gg 1$ , i.e, a probe which is heavy compared to the temperature. In this case the first term in (62) is just the particle freely traversing the thermal space. The result in (63) is suppressed by a factor of  $e^{-\beta w_1/2}$ . This reproduces the suppression one would expect from propagation throughout the Euclidean instanton, where the disconnected initial time surfaces correspond to the constant time slices at  $\tau = 0, \beta/2$ .

## 6 Black holes and random operators

We conclude with a more speculative section. The main conjecture in this paper is that low energy gravity modes are well approximated by random operators on any state which appears as a black hole for an outside observer, and that using this approximation we can find a simple criteria for when an Einstein-Rosen throat is semi-classical. Since most discussions of the information paradox [24] or firewalls [25] are phrased in terms of such probes, it will be interesting to explore the consequences of their conjectured randomness in this regard. We briefly discuss some issues in this direction, which are the robust analytic structure of such operators and the problem of “seeing behind the horizon”, and the possible emergence of a new perturbation theory for random operators, with some speculations on instances where this perturbation theory exhibits unusual behaviour for probes which cross close to the horizon.

### 6.1 The analytic structure of random operators in pure states

We modelled supergravity probes as random because they are insensitive to the details of the black hole state. One might think this would make it difficult to get any sensible coherent behaviour behind the horizon. However, precisely the opposite is true. Since the correlations don’t depend on the state, we get an effective coarse-graining over the possible microstates of the black hole. As a result, the analytic structure of random operators is more robust than for more structured operators, allowing us to carry out the continuation to the “other side” for a large class of ensembles or states on  $\mathcal{H}_R$ .

To see this recall equations (38), (39) and (42) above, which gave the same single sided two point function for any generic single sided state or density matrix (up to  $1/N^2$  corrections),<sup>17</sup> so the single sided correlation functions do not depend on the particular state we consider, they all mimic the thermal density matrix. This means that in any such state or density matrix we can continue to the second copy in exactly the same way. This result relies heavily on the random matrix aspect of the probes, and it seems unlikely to us to hold without some assumption on the spectral properties of the probe operators.

Thus, if we are in any of these density matrices, the value of the single-sided two point function is as in the thermal ensemble. Just as for the thermal ensemble, we can then

---

<sup>17</sup>By generic state we mean a state that we do not arrange using the spectral data of  $\mathcal{O}_R$ . For example we do not allow ourselves to find the eigenvectors of  $\mathcal{O}_R$  and then choose one of them as our state.

define the action of operators on a second copy of the Hilbert space  $\mathcal{H}_L$  as in section 3.3.1. The small differences in the single-sided correlations will grow exponentially under this continuation, but it was noted in [26] that for a typical pure state these differences are small enough that they are still exponentially suppressed after analytic continuation to  $t' = t + i\beta/2$ .

This result may seem to be in tension with our previous result that the two-sided correlators in the typical state on  $\mathcal{H}_L \otimes \mathcal{H}_R$  are very different from the one-sided correlators. The point here is that we are considering the analytic continuation of the one-sided correlators, based on the fact that these do not contain information on the actual state or density matrix considered. These thus do not correspond to the actual two-sided correlators in a given typical state on  $\mathcal{H}_L \otimes \mathcal{H}_R$ .

### 6.1.1 Relation to the Papadodimas-Raju construction

Our conclusion that for random operators, the operators on the “other side” can be defined independent of the choice of state or ensemble is rather similar to the premise underlying the work of Papadodimas and Raju [27], that mirror operators can be defined for any pure state of the black hole. Our discussion of random operators provide a different perspective on when such an analytic continuation should be possible, which can be contrasted to Papadodimas and Raju’s discussion in terms of large-N factorization and generalized free fields.

One of the differences is that the construction of the mirror operators in [27] was highly state-dependent, whereas our basic point was that it was the state-independence of the correlators which makes the analytic continuation robust. Our goals are different: we have discussed the analytic continuation of the operator correlation functions, whereas [27] aimed to explicitly construct an operator acting on the microscopic Hilbert space whose correlation functions reproduced these analytically continued correlators.

It would be interesting to compare the two approaches further, and in particular whether the two approaches converge, up to small corrections, if one makes further spectral assumptions about the operators within the approach of [27]. In any case, we believe that both can be a good starting point for exploring perturbation theory in the black hole background, and deviations from it.

## 6.2 Emergence of perturbation theory and modified horizon dynamics

In the standard description of low energy gravity probes, one quantizes bulk fields in the eternal black hole background and interprets their quantum state in terms of creation and annihilation operators (at least perturbatively and in the absence of backreaction). This is certainly the case in Hawking’s original calculation [28] and it is also the point made in the reconstruction of bulk operators from CFT ones in [29, 27]. This picture becomes more involved when trying to embed this perturbative Hilbert space structure in the dual field theory Hilbert space [27, 30].

In this standard approach, there is a well defined bulk perturbative expansion in terms of Feynman diagrams. In this subsection we ask whether there is a corresponding notion of perturbation theory in our random matrix description.

The relation of the gaussian ensemble to free fields was discussed in section 3.1. The

gaussian matrix models gave ensemble average two-point functions

$$\mathbf{E}(M_{ij}M_{kl}^*) = \frac{\delta_{ik}\delta_{jl}}{\gamma\Delta_{ij}}, \quad (64)$$

so insertions of  $\mathcal{O}$  are pairwise contracted, giving rise to a Wick-like expansion of correlators (after summing over the indices).

We can speculate on how this will be extended to an interacting picture. An interaction term in the bulk, for example an  $n$ -particle vertex, means that we can tie  $n$  propagators together. In the eikonal approximation, it means that  $n$  insertions of  $\mathcal{O}$  (or it's conjugate) can be contracted. In the matrix model, we could model this by including higher order polynomials in our ensemble distributions (10) and (32). Expanding the exponents of these polynomials will give rise to perturbative insertions of these vertices in our correlator ensemble averages  $\mathbf{E}(M_{ij} \dots M_{kl}^*)$ , which would match with the spacetime contraction. All of this, of course, applies to interactions close to or inside the black hole, within the part of the bulk correlation functions that our random matrix is supposed to model.

This approach would correspond to diagonalizing the full Hamiltonian in the interacting theory. An alternative approach would be to write a gaussian matrix ensemble based on diagonalizing the free Hamiltonian, and write the full Hamiltonian as  $H = H_0 + P(\mathcal{O}, \mathcal{O}^\dagger)$ , for some polynomial  $P$  in the operator  $\mathcal{O}$  and its conjugate. The interactions would then enter by including  $P$  perturbatively in the time evolution.

But before developing perturbation theory, we can see that even at the "free field" level, the random operator Wick-like scheme has peculiar features. The reason for this is that the random matrix contraction applies to operator matrix elements rather than to the operators themselves. Consider a four-point function of the form  $\langle \psi_\beta | \mathcal{O}_R^\dagger \mathcal{O}_R^\dagger \mathcal{O}_R \mathcal{O}_R | \psi_\beta \rangle$ . In standard bulk perturbation theory, we would contract all possible pairings of  $\mathcal{O}_R$  and  $\mathcal{O}_R^\dagger$ . Each such contraction would contribute equally to the full bulk correlation function, up to some time dependence, i.e., none of them would be suppressed with the dimension of  $\mathcal{H}_E$ . We want to show this is not the case within the perturbative expansion using random matrices.

In the random matrix formalism, the objects we compute are matrix element ensemble averages such as

$$C_1 = \mathbf{E}(M_2^\dagger M_1^\dagger M_2 M_1), \quad C_2 = \mathbf{E}(M_2^\dagger M_2 M_1^\dagger M_1), \quad (65)$$

where we have put subscripts 1 and 2 on the matrices  $M$  in the order of their appearance in the correlator; all  $M$ 's refer to matrix elements of the same operator. The  $C_2$  correlator allows contractions  $M_1 - M_1^\dagger$ ,  $M_2 - M_2^\dagger$ , which gives

$$\mathbf{E}(M_{k_1 i}^* M_{k_1 k_2}) \mathbf{E}(M_{k_3 k_2}^* M_{k_3 i}) = \delta_{k_1 k_1} \delta_{i k_2} \delta_{k_3 k_3} \delta_{k_2 i} = d_E^3. \quad (66)$$

and  $M_1 - M_2^\dagger$ ,  $M_1^\dagger - M_2$ , which similarly gives

$$\mathbf{E}(M_{k_1 i}^* M_{k_3 i}) \mathbf{E}(M_{k_3 k_2}^* M_{k_1 k_2}) = \delta_{k_1 k_3} \delta_{i i} \delta_{k_1 k_3} \delta_{k_2 k_2} = d_E^3. \quad (67)$$

On the other hand, for the  $C_1$  correlator, while the contraction  $M_1 - M_2^\dagger$ ,  $M_1^\dagger - M_2$  gives

$$\mathbf{E}(M_{k_1 i}^* M_{k_3 i}) \mathbf{E}(M_{k_2 k_1}^* M_{k_2 k_3}) = \delta_{k_1 k_3} \delta_{i i} \delta_{k_2 k_2} \delta_{k_1 k_3} = d_E^3 \quad (68)$$

with the same scaling as before, the contraction  $M_1 - M_1^\dagger$ ,  $M_2 - M_2^\dagger$  involves

$$\mathbf{E}(M_{k_1 i}^* M_{k_2 k_3}) \mathbf{E}(M_{k_2 k_1}^* M_{k_3 i}) = \delta_{k_1 k_2} \delta_{i k_3} \delta_{k_2 k_3} \delta_{k_1 i} = d_E, \quad (69)$$

and is suppressed by a factor of  $e^{-2S}$ . This can be understood by standard planarity arguments: the last combination involves a non-planar contraction of the matrix indices, because our Wick-like contractions in  $C_1$  cross over in this case.

From the bulk point of view, the correlators we are averaging describe the insertion and extraction of two particles. In the eikonal approximation, we can view each contraction  $\mathbf{E}(M_{ij}M_{kl}^*)$  as approximating a spacetime geodesic from the insertion point of  $\mathcal{O}_R$  to the insertion point of  $\mathcal{O}_R^\dagger$ . Thus, in the bulk the distinction between the suppressed term and the other terms is that the associated bulk geodesics are crossing in the near-horizon region<sup>18</sup> in the suppressed term (one comes out of in between the times when the other went in and came out), while the other terms involve non-crossing bulk geodesics.

Thus, our random operator model seems to say that not all combinations of geodesics from entry to exit points contribute equally. This may represent a potentially serious challenge for our random matrix ensemble description. However, there are known to be subtleties in determining which spacetime geodesics actually contribute to correlations and which do not [32]. Alternatively, it may be a signal that our random matrix description is capturing some back-reaction effect along the lines of [33].

The discussion above is very preliminary in several ways. First, it just involves the rules of free field theory, and no local interactions were included. Second, there is no notion of radial direction (or any bulk dimension for that matter) in our formalism, which is known to be a difficult problem in AdS/CFT. Finally, to make the range of validity of our approach more precise we need to describe more accurately the transition from the structured part of the operator to its unstructured part using the spectral properties of the underlying OPEs in the exact dual CFT description. We leave all this for future work.

## Acknowledgements

We would like to thank Ofer Aharony, José Barbón, Nima Lashkari, Amir Zait and Ofer Zeitouni for interesting discussions. SFR and JS would like to thank the organisers of the COST-WIS workshop on *Black holes and quantum information* for hospitality during the last stages of this project. The work of JS was partially supported by the Engineering and Physical Sciences Research Council (EPSRC) [grant number EP/G007985/1] and the Science and Technology Facilities Council (STFC) [grant number ST/J000329/1]. The work of SFR is supported in part by STFC. The work of MB is supported by an ISF center of excellence grant, GIF and the Minerva foundation. VB is partly supported by DOE grant DE-FG02-05ER-41367.

## A Corrections to correlators

We have emphasized before that single sided correlators are the same for any state or ensemble, which is peaked around a given energy, in particular in the microcanonical, canonical ensembles and some of the pure states. The discussion in section 6, on “seeing behind the horizon”, hinged on this assertion. In this appendix we will quantify the differences between these expressions, i.e., examine the size of the corrections when going from (38) and the

---

<sup>18</sup>We refer to the near horizon region because our random matrix description is supposed to capture the behaviour of our low energy gravity probes close to the horizon, as explained in section 2.



analogue formula for the microcanonical ensemble, which is

$$\int dE_k \frac{e^{S(E_k)+i(E_i-E_k)t}}{\hat{\gamma}(e^{S(E_i)} + e^{S(E_k)})P(\alpha, E_k - E_i)} \quad (70)$$

to (39).

To go from (70) to (39) we just expand  $S(E_i) = S(E_k) - \beta(E_k)\Delta + O(\Delta^2/\textit{extensive})$  where  $\Delta = E_k - E_i$ . We then obtain expression (39), up to the  $O(1/\textit{extensive})$  corrections, where by  $1/\textit{extensive}$  correction we mean corrections that scale like  $1/\textit{energy}$  or  $1/\textit{entropy}$ . To go from the canonical ensemble expression (38) to (39) we carry out the same expansion and change the integration variables to  $E_i$  and  $\Delta$ . This part gives rise to  $1/\textit{extensive}$  corrections as before. The integral over  $E_i$  localizes around the average energy in the ensemble  $\bar{E}$ , up to  $E_i - \bar{E} \sim \sqrt{\textit{extensive}}$ . We are then justified to expand  $\alpha(E_i)$  around  $\bar{E}$  as  $\alpha(\bar{E}) + (1/\textit{extensive})(E_i - \bar{E}) + ..$  which gives rise to additional  $1/\textit{extensive}$  corrections. Hence, the canonical and microcanonical correlators are the same up to  $1/\textit{extensive}$  corrections. In addition, there could be  $1/N^2$  corrections to these formulas, which could be larger than the  $1/\textit{extensive}$  corrections identified in this appendix.

## References

- [1] J. M. Maldacena, “Eternal black holes in anti-de Sitter,” JHEP **0304** (2003) 021 [hep-th/0106112].
- [2] M. Van Raamsdonk, “Building up spacetime with quantum entanglement,” Gen. Rel. Grav. **42** (2010) 2323 [Int. J. Mod. Phys. D **19** (2010) 2429] [arXiv:1005.3035 [hep-th]].
- [3] J. Maldacena and L. Susskind, “Cool horizons for entangled black holes,” arXiv:1306.0533 [hep-th].
- [4] S. Ryu and T. Takayanagi, “Holographic derivation of entanglement entropy from AdS/CFT,” Phys. Rev. Lett. **96** (2006) 181602 [hep-th/0603001].
- [5] V. Balasubramanian, B. D. Chowdhury, B. Czech, J. de Boer and M. P. Heller, “A hole-ographic spacetime,” Phys. Rev. D **89**, 086004 (2014) [arXiv:1310.4204 [hep-th]].
- [6] R. C. Myers, J. Rao and S. Sugishita, “Holographic Holes in Higher Dimensions,” arXiv:1403.3416 [hep-th].
- [7] B. Swingle, “Constructing holographic spacetimes using entanglement renormalization,” arXiv:1209.3304 [hep-th].
- [8] D. Marolf and J. Polchinski, “Gauge/Gravity Duality and the Black Hole Interior,” Phys. Rev. Lett. **111** (2013) 171301 [arXiv:1307.4706 [hep-th]].
- [9] S. H. Shenker and D. Stanford, “Black holes and the butterfly effect,” JHEP **1403** (2014) 067 [arXiv:1306.0622 [hep-th]].
- [10] S. H. Shenker and D. Stanford, “Multiple Shocks,” arXiv:1312.3296 [hep-th].
- [11] V. Balasubramanian and S. F. Ross, “Holographic particle detection,” Phys. Rev. D **61**, 044007 (2000) [hep-th/9906226].

- [12] V. Balasubramanian, V. Jejjala and J. Simon, “The Library of Babel,” *Int. J. Mod. Phys. D* **14** (2005) 2181 [hep-th/0505123]; V. Balasubramanian, J. de Boer, V. Jejjala and J. Simon, “The Library of Babel: On the origin of gravitational thermodynamics,” *JHEP* **0512** (2005) 006 [hep-th/0508023].
- [13] V. Balasubramanian, P. Kraus and M. Shigemori, “Massless black holes and black rings as effective geometries of the D1-D5 system,” *Class. Quant. Grav.* **22**, 4803 (2005) [hep-th/0508110].
- [14] N. Lashkari and J. Simon, “From state distinguishability to effective bulk locality,” arXiv:1402.4829 [hep-th].
- [15] S. G. Avery and B. D. Chowdhury, “No Holography for Eternal AdS Black Holes,” arXiv:1312.3346 [hep-th].
- [16] S. D. Mathur, “What is the dual of two entangled CFTs?,” arXiv:1402.6378 [hep-th].
- [17] L. Susskind, “Computational Complexity and Black Hole Horizons,” arXiv:1402.5674 [hep-th].
- [18] L. Susskind, “Addendum to Computational Complexity and Black Hole Horizons,” arXiv:1403.5695 [hep-th].
- [19] O. Lunin and S. D. Mathur, “AdS / CFT duality and the black hole information paradox,” *Nucl. Phys. B* **623** (2002) 342 [hep-th/0109154].
- [20] I. Bena and N. P. Warner, “One ring to rule them all ... and in the darkness bind them?,” *Adv. Theor. Math. Phys.* **9** (2005) 667 [hep-th/0408106].
- [21] J.M. Deutsch, *Phys. Rev. A* **43** (1991) 2046.
- [22] M. Srednicki, *Phys. Rev. E* **50** (1994) 888 [arXiv:cond-mat/9406056]; *J. Phys. A* **29** (1996) L75 [arXiv:cond-mat/96]; *J. Phys. A* **32** (1999) 1163 [arXiv:cond-mat/9809360];
- [23] T. Banks, M. R. Douglas, G. T. Horowitz and E. J. Martinec, “AdS dynamics from conformal field theory,” hep-th/9808016.
- [24] S. D. Mathur, “The Information paradox: A Pedagogical introduction,” *Class. Quant. Grav.* **26**, 224001 (2009) [arXiv:0909.1038 [hep-th]]; S. D. Mathur, “What the information paradox is not,” arXiv:1108.0302 [hep-th].
- [25] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, “Black Holes: Complementarity or Firewalls?,” *JHEP* **1302**, 062 (2013) [arXiv:1207.3123 [hep-th]]; cf. S. L. Braunstein, “Black hole entropy as entropy of entanglement, or it’s curtains for the equivalence principle,” [arXiv:0907.1190v1 [quant-ph]] published as S. L. Braunstein, S. Pirandola and K. Zyczkowski, “Better Late than Never: Information Retrieval from Black Holes,” *Physical Review Letters* **110**, 101301 (2013) for a similar prediction from different assumptions.
- [26] V. Balasubramanian, B. Czech, V. E. Hubeny, K. Larjo, M. Rangamani and J. Simon, “Typicality versus thermality: An Analytic distinction,” *Gen. Rel. Grav.* **40** (2008) 1863 [hep-th/0701122].

- [27] K. Papadodimas and S. Raju, “The Black Hole Interior in AdS/CFT and the Information Paradox,” *Phys. Rev. Lett.* **112**, 051301 (2014) [arXiv:1310.6334 [hep-th]].  
K. Papadodimas and S. Raju, “State-Dependent Bulk-Boundary Maps and Black Hole Complementarity,” arXiv:1310.6335 [hep-th].
- [28] S. W. Hawking, “Particle Creation by Black Holes,” *Commun. Math. Phys.* **43** (1975) 199 [Erratum-ibid. **46** (1976) 206].
- [29] A. Hamilton, D. N. Kabat, G. Lifschytz and D. A. Lowe, “Local bulk operators in AdS/CFT: A Boundary view of horizons and locality,” *Phys. Rev. D* **73** (2006) 086003 [hep-th/0506118].  
A. Hamilton, D. N. Kabat, G. Lifschytz and D. A. Lowe, “Holographic representation of local bulk operators,” *Phys. Rev. D* **74** (2006) 066009 [hep-th/0606141].  
D. Kabat, G. Lifschytz and D. A. Lowe, “Constructing local bulk observables in interacting AdS/CFT,” *Phys. Rev. D* **83** (2011) 106009 [arXiv:1102.2910 [hep-th]].
- [30] E. Verlinde and H. Verlinde, “Behind the Horizon in AdS/CFT,” arXiv:1311.1137 [hep-th].
- [31] S. Lloyd, “Pure state quantum statistical mechanics and black holes,” [arXiv:1307.0378].
- [32] P. Kraus, H. Ooguri and S. Shenker, “Inside the horizon with AdS / CFT,” *Phys. Rev. D* **67** (2003) 124022 [hep-th/0212277].
- [33] G. ’t Hooft, “The Scattering matrix approach for the quantum black hole: An Overview,” *Int. J. Mod. Phys. A* **11** (1996) 4623 [gr-qc/9607022].